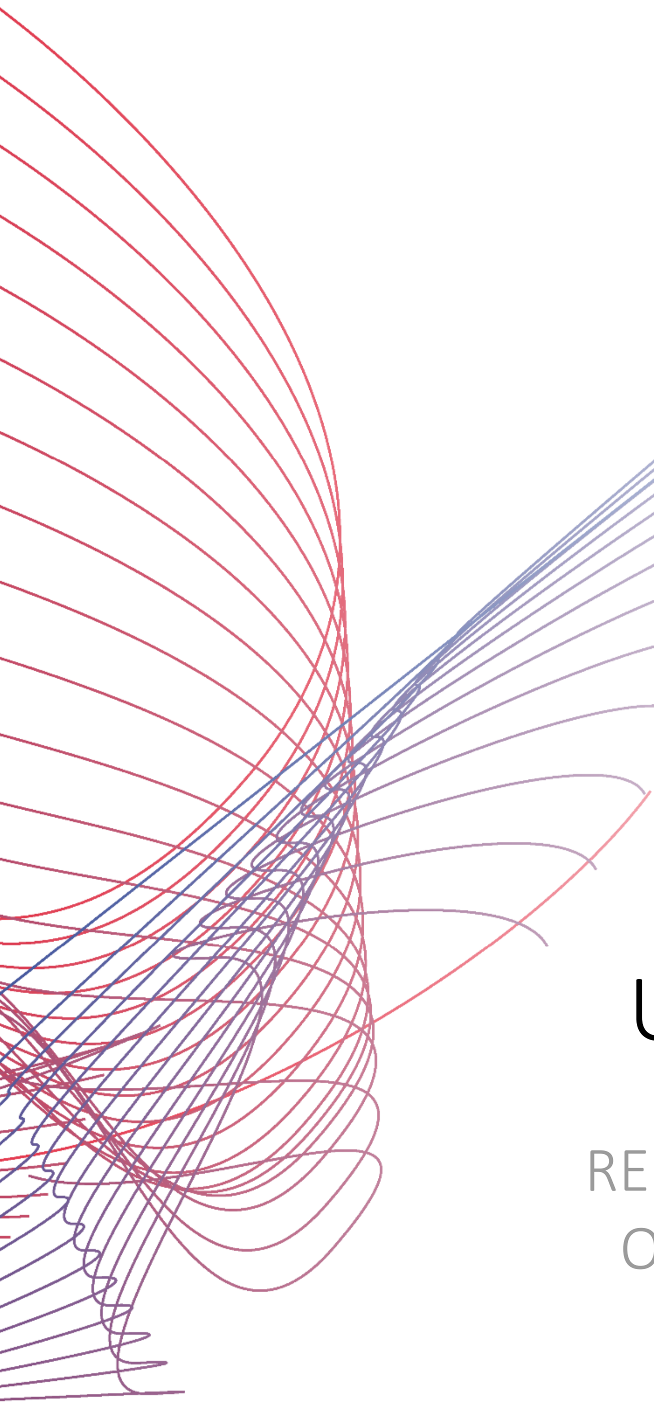
en
cs de
sk es

TECH pedia



USER IDENTIFICATION

RENATA RYBÁROVÁ, JURAJ KAČUR,
ONDERJ LÁBAJ, GREGOR ROZINAJ

Title:	User identification
Author:	Renata Rybárová, Juraj Kačur, Onderj Lábaj, Gregor Rozinaj
Published by:	Czech Technical University of Prague
	Faculty of electrical engineering
Contact address:	Technicka 2, Prague 6, Czech Republic
Phone Number:	+420 224352084
Print:	(only electronic form)
Number of pages:	43
Edition:	1st Edition, 2017
ISBN	978-80-01-06237-1

TechPedia
European Virtual Learning Platform for
Electrical and Information Engineering

http://www.techpedia.eu

This project has been funded with support from the European Commission.
This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Explanatory Notes

$E=m\cdot c^2$	Definition
⚑	Interesting
i	Note
🧪	Example
🧩	Summary
+	Advantage
−	Disadvantage

ANNOTATION

User identification, authorization and authentication assure that the system is only used by certain users and only commands which are properly authorized are performed. The speaker identification tends to provide basic identification of the possible users located in the system installation area. This would be suitable for identification tasks, such as loading personal profile. The face detection approach aims to provide more reliable user identification based on users' faces which contain far more characteristics that can be parameterized in comparison to the voice identification approach. Additionally, the 3D face recognition further extends the possibilities of feature extraction in order to more precisely identify particular persons and can be thus used for the highest level authentication (and authorization) for the most demanding applications (e.g. bank account login, etc.).

OBJECTIVES

The main goal of the module is to introduce a student to the fundamental of user identification, authentication and authorization processes. The student is clearly acquainted with the base principles of speaker identification, user identification based on 2D and 3D techniques of face recognition, authentication methods and user authorization.

LITERATURE

[1] Abate, Andrea F.; Nappi, Michele; Riccio, Daniel; Sabatino, Gabriele. 2D and 3D face recognition: A survey In: Pattern Recognition Letters, Volume 28, Issue 14, 15 October 2007, Pages 1885–1906. available at www.sciencedirect.com.

[2] T. Kinnunen, H. Li, An overview of text-independent speaker recognition: from features to supervectors, Speech communication, Vol. 52, pp. 12-40, Elsevier, 2010

[3] Probst, Michael; Schumann, Sebastian; Rozinaj, Gregor; Minarik, Ivan; Rybárová, Renata; Oravec, Miloš. EVALUATION: Final Multimodal Interface for User/Group-Aware Personalisation, Deliverable 5.5.1, available at http://www.hbb-next.eu/index.php/documents, Decmber 2013.

[4] Bán, Jozef; Féder, Matej; Oravec, Miloš; Pavlovičová, Jarmila. Face Recognition of Images Corrupted by Transmission Errors. In: Redžúr 2012: proceedings; 6th International Workshop on Multimedia and Signal Processing. April 11, 2012, Vienna, Austria. Bratislava: Nakladateľstvo STU, 2012. pp. 15-18, ISBN 978-80-227-3686-2

[5] Rozinaj, Gregor; Minarik, Ivan; Rybárová, Renata; Pavlovičová, Jarmila; Mármol, Félix Gómez; Tormo, Ginés Dólera, Gülbahar, Mark; Schumann, Sebastian. DESIGN AND PROTOCOL: Final User ID, Profile, Application Reputation Framework, Deliverable 3.4.1, available at http://www.hbb-next.eu/index.php/documents, Decmber 2013.

[6] Schneier, Bruce. Sensible Authentication, ACM Queue 1, Volume 1 Issue 10, February 2004. Pages 74.

[7] McCue, A. Is Your Cat a Target for Password-Stealing Hackers?, Silicon.com, 11 August 2004.

[8] Haskett, J.A., Pass-Algorithms: A User Validation Scheme Based on Knowledge of Secret Algorithms In Communications of the ACM 27, 1984.

[9] Madigan, A. Picture Memory - Memory and Cognition: Essays in Honour of Allan Paivio Erlbaum, 1983.

[10] Cranor, L.F.; Garfinkel, S. Security and Usability, O'Reilly, August 2005. ISBN 0-596-00827-9.

[11] Vacca, J.R. Computer and Information Security Handbook, Morgan Kaufmann, 2009. ISBN 978-0-12-374354-1.

[12] Gattiker, U. E. The Information Security Dictionary, KLUWER ACADEMICPUBLISHERS, 2004. ISBN 1-4020-7927-3.

Index

1 User identification

User Identification is one of the key features assuring that the system or application only performs commands which are properly authorized. The most widely used authentication type is password, but with development of information technologies and security protection algorithms, systems and applications start to use authentication based on the biometric factors.

+

Using biometrics effectively eliminate possible risks associated with less advanced technologies that are based on what a person have or know rather than whom a person really is [1]. It is a very attractive and popular technology, because it can be integrated into any application or system requiring security or access control.

The speaker identification tends to provide basic identification of the possible users located in the system installation area. The face detection approach aims to provide more reliable user identification based on users' faces which contain far more characteristics that can be parameterized in comparison to the voice identification approach. Additionally, the 3D face recognition further extends the possibilities of feature extraction in order to more precisely identify particular persons and can be thus used for the highest level authentication (and authorization) for the most demanding applications (e.g. bank account login, etc.). For security reasons, 3D face recognition authentication can be enhanced with e.g. eye movement tracking or iris recognition. This approach may simulate multi-factor authentication (login plus token) necessary for the highest level authentication.

−

However, biometrics have also drawbacks. Iris recognition is extremely accurate, but expensive to implement and not very accepted by people. Fingerprints are reliable and non-intrusive, but not suitable for non-collaborative individuals. On the contrary, face recognition seems to be a good compromise between reliability and social acceptance [1].

2 Speaker identification

2.1 Speaker identification overview

Speaker identification is a part of a broader concept known as speaker recognition. It encompasses two important and in a way similar but still different tasks, namely: speaker identification and speaker verification. The first one basically points to an assignment of automatically deciding who the tested voice sample belongs to from a set of users stored in a database during an enrolment (training) phase.

Optionally if the confidence of a final decision is too low nobody is recognized. This task is often referred to as a closed group problem, as there is fixed set of user who may be recognized.

On the other hand verification process evaluates whether the tested individual is the one who he or she claims to be.

As there are many users (possibly 7 billions of living people) it is impossible to have measured characteristic or model for everyone. This task is therefore referred to as open group problem. In this situation model of a general speaker is vital to determine proper acceptance/rejection thresholds.

The speaker recognition is quite a difficult problem because of many reasons mentioned later in the text and has been under serious scientific investigation for over 40 years by many research teams. As there are emerging new and accessible technologies it is finding rapidly grooving application in many areas, just to mention few of them:

- forensic science
- natural and non-invasive method for secure access and protection of data and service
- automatic indexing of speech and audio recordings stored in databases
- range of applications for game industry
- aids for disabled people

Because of the wide spectrum of problems that must be tackled with there are many solutions and common techniques related to speaker identification problem. Those can be classified into 3 main groups:

- **speech features** - proper for speaker recognition or speaker identification task
- **feature normalization/ model compensation techniques**- that aims to supress session variability

- **classification and decision taking algorithms**- decide based on features and models who represents the closes match to the unknown sample

Finally the task of speaker identification is divided into two major groups i.e. text dependent and text independent problem. In the first group the identification process doesn’t assume any specific text while in the later class the systems require precise text to be uttered. Obviously text dependent system reaches higher accuracies.

2.2 Properties of speech signals

Genuine speech signals are created and produced by human beings; more precisely by their vocal apparatus and their brains which are unique to every individual. Both phases naturally leave their marks in the audible signal, and thus the speech can be regarded as a biometric signal.

Even though the main goal of speech signals is to convey lexical information it contains. Except the lexical part which is roughly given by the sequence of different positions of vocal organs it contains biometric information about any speaker represented mainly by different shapes, sizes, weights and toughness of vocal organs, actual mood of a person (intonation, speech pace, stress, etc.), and social background of a person (dialect, vocabulary, etc.).

However these different pieces of information are encoded in the speech signal by a difficult transformation which is believed to be irreversible and not known. Thus it is a difficult problem to extract just the information which is needed for a particular task (lexical, identification, mood, health state, …). Furthermore speech exhibits great within speaker variability given by individual's current mood, health and physical state or other conditions. Finally the acoustical form of a speech signal can be seriously altered by differences in recording devices, room where it was recorded and present background noise.

The modifications of speech that are not related to a speaker (devices, room, etc.) are called session variability. This aspect causes major problems and must be dealt with in situation where the enrolment conditions do not match the actual deployment one.

2.3 Feature extraction

Because of the variable speech characteristics and many adverse conditions mentioned in the previous text, there have been many extraction techniques invented through the time. Basically, a good speech feature must be:

- discriminative
- robust against different background noises
- insensitive to changes caused by recording devices and places of deployment
- supress speaker variability
- easy to compute and process

As there are many different speaker specific features that have different physical meanings we distinguish 3 sorts of features (from the speaker recognition point of view):

- Acoustic
- Prosodic
- Higher level

At the acoustic level short time features are gathered that are related to physical characteristic of vocal apparatus. These methods mainly represent modified spectral (envelop) shapes extracted from intervals ranging form 10ms to 30ms. Further, they apply different psychoacoustic principles of human haring system to increase their robustness. At present the most common are *Mel frequency cepstral coefficients* (**MFCC**), *Perceptual Linear Prediction* **(PLP)**, or *Cepstral Linear Prediction coefficients* **(CLPC)** features. MFCC and PLP try to capture modified spectral envelopes following some psychoacoustic principles like critical bands, human perception of frequencies, equal loudness curve, conversion of intensities to loudness, etc. As they are able to extract spectral envelopes they preserve and emphasise the location, widths and shapes of formant frequencies that are vital for the perception of differences among sounds. Thus they are very important for the speech recognition systems. Even thou they still play major roles for the speaker recognition problem as well. It can be explained so that they are able to capture slight differences in locations and shapes of formant frequencies that vary from person to person as present in particular phones. CLPC features are based on modelling the speech production mechanism instead of the hearing and perception process. Finally to encompass dynamic of acoustical features in the time, differential and acceleration coefficients may be derived as well. As they cover longer time intervals they may detect differences in co-articulation that are specific to a particular speaker.

The prosodic level focuses more on the style of speaking, mood of a speaker, specific speaking habits, physical and health conditions, etc. Obviously this information is located and can be extracted using only longer time intervals spreading several seconds of speech. The most favourite features for this level are:

rhythm, speech dynamics, pace of speaking, modulation of fundamental frequency, sort of pauses made while speaking, etc. However these features are more difficult to measure and qualify as those on the acoustic level. Thus there are several methods to extract and evaluate them over the proper time interval. The most common approaches are the autocorrelation function, *Average Magnitude Difference Function* **(AMDF)** function, inverse filtering for the fundamental frequency detection, energy for the speech dynamics and so on. However there are many modifications to both autocorrelation and AMDF.

2.4 Classification/decision taking algorithm

After the feature extraction and possible normalization/ compensation phase (will be presented in the next section) a classification must be used to decide which user (features or models) are the closes to the unknown one. Still it is possible to reject any if the recorded match/ confidence is too low. There are several successful classification techniques that differ in their complexity, way how they work and what they assume upon the processed data. These methods are classified into several major categories each with its pros and cons as follows:

- **Non-parametric methods**- do not impose any restriction on the data, thus they do not use any model to describe the space. Its main representative is the *K-nearest neighbour method* **(KNN)** KNN finds k closest vectors to the unknown one and based on them use some criteria to decide about the result.

- **Parameter based methods**- assume some structure of the feature space and model it by some parameters. The most successful and frequently used is the mixture of Gaussian distributions. This model is called *Gaussian mixture model* **(GMM)**. In the lack of data and using proper models that match the space, these methods are superior to the non –parametric one.

- **Discriminative methods** – try to model/ divide the feature space so that the classification error is as small as possible. It is quite straightforward to do this on the training part of the database but it is trickier to do it (only estimate) for unseen data. In such situation good generalization ability is required (low error for unseen samples). The main representatives of this group are *neural networks* (**NN**) and *support vector machines* (**SVM**). Both NN and SVM can in certain conditions provide excellent results.

- **Generative methods** - aim to describe the space as close as possible and do not attempt to separate individual users. If the models perfectly match the data (in reality they don’t) it is possible to construct an optimal classifier that reach the minimal possible cost which was derived by Bayes, thus it is referred to as Bayes classification. Again, the most successful model is the above mentioned GMM.

It was discovered that it is beneficial to use general models of universal speakers that have been created by training samples from many speakers.

2.5 Environment compensation

To decrease the session variability caused by different training- testing conditions (background noises, different acoustic parameters of recording devices and rooms) there have been invented and employed several concepts that do this. The most basic methods uniformly normalize signal's dynamic by manipulating the overall power or equalize the power of each frequency band of an averaged speech spectrum which is usually done by cepstral mean subtraction. Further it is possible to use fixed filtering techniques that emphasise a general speech signal like amplifying the speech modulation spectrum or *relative spectral analysis* **(RASTA)** filtering. More sophisticated methods try to find optimal transformations mapping enrolment features to features observed in the deployment environment (so called feature mapping methods) or to transform whole models of speakers to match the model of the employment environment (it is called speaker model synthesis). However, these methods are based on higher mathematics, and adapt their behaviour with the incoming data. Thus if the employment environment is changing so does the optimal mapping.

Another less sophisticated but some time useful solution is to have pre-recorded speech samples (features or models) in different conditions and prior to the recognition detect the proper one. Then use the best match environment for a particular recording. It is obvious that the best results are observed when there is a match between training and testing environments.

To get a more detailed overview on the topic of speaker recognition please sees e.g. [2].

3 Face recognition

Face become the most attractive biometric and face recognition systems for personal identification are increasingly used in a wide range of applications. Development of recognition algorithms and methods made possible usage of identification and verification systems in commercial field. However these systems do not achieve comparable recognition rates under uncontrolled and unconstrained conditions. Face recognition under these conditions is still a challenging problem in spite of recent advances in well-constrained face recognition.

Biometric systems for personal identification, which are developed by several vendors, achieve very high face recognition accuracy. The most of these applications require [3]:

- recognition systems which can recognize several faces from one video frame or one image
- high recognition rate
- illumination invariance
- stability under changing of face expressions and pose
- recognition in real time, etc.

So we can see there is several factors which can affect performance and accuracy of system for face recognition [1]:

- **Illumination** variations due to skin reflectance properties and due to the internal camera control. Several 2D methods do well in recognition tasks only under moderate illumination variation.
- **Pose changes** affect the authentication process, because they introduce object deformations. Detection methods should solve the problem considering various viewing angles when object is positioned (e.g. view from security cameras). On the other hand the algorithms are relatively robust to facial expression (except some extreme expressions like scream).
- The **time delay** is also important factor, considering the face changes over time, in a nonlinear way over long periods (age variations). In general this problem is harder to solve with respect to the others.

3.1 Face recognition methods

Face recognition systems fall into two categories: verification and identification.

Face verification is a 1:1 match. In this process face image whose identity is being claimed is compared to a template face images.

On the contrary, face identification is a 1:N problem. Face image is compared to all image templates in a face database to determine the identity of the query face.

In case we do not know if the tested face is in the system database, process is as following. The query face image is also compared against all the face images in the database, evaluating probability for each one. All these probabilities are numerically ranked: the highest value is first. In case probability is higher than given threshold, system notifies us about the result [1].

The selected basic 2D face recognition methods:

- Linear/nonlinear projection methods
  - *Principal Component Analysis* (**PCA**) - the method based on PCA is named eigenface. The major idea of PCA is to decompose a data space into a linear combination of a small collection of bases, which are pair wise orthogonal and which capture the directions of maximum variance in the training set [4].
  - *Kernel Principal Component Analysis* (**KPCA**) - is a method of non-linear feature extraction. The KPCA can extract the feature set which is more suitable in categorization than the conventional PCA. The KPCA has been widely used in the case of face recognition with face expression and under varying illumination [4].
  - *Linear Discriminant Analysis* (**LDA**) - has been proposed as a better alternative to the PCA. It provides discrimination among the classes, while the PCA deals with the input data in their entirety, without paying any attention for the underlying structure. Indeed the main aim of the LDA consists in finding a base of vectors providing the best discrimination among the classes, trying to maximize the between-class differences, minimizing the within-class ones [1].
  - *Discriminant Common Vectors* (**DCV**) - the main idea of the DCV consists in collecting the similarities among the elements in the same class dropping their dissimilarities [1].
- The neural networks – is nonlinear solution, is used also in other pattern recognition problems. The advantage of neural classifiers over linear ones is that they can reduce misclassifications among the neighborhood classes. The basic idea is to consider a net with a neuron for every pixel in the image. Nevertheless, because of the pattern dimensions neural networks are not directly trained with the input images, but the application of a dimensionality reduction technique is used before training [1].

- Fractal and *iterated function systems (IFS)* – IFS theory has been mainly developed in the area of image coding and lately has been extended to the image indexing. Fractal code of an image is invariant with respect to a wide set of global transformations, such as rotations, contrast scaling, etc. IFS fractal of a face image is used for training a neural networks, where is used as clasifier [1].

3.2 Feature extraction

Some facial recognition algorithms are based on features extracted from an image of the subject's face – on facial features. For example, an algorithm may analyze the relative position, size, and/or shape of the eyes, nose, mouth, cheekbones, and jaw. These features are then used during search in group of images for matching features. Other algorithms normalize a gallery of face images and then compress the face data, saving only the data in the image that is useful for face recognition. A tested image is then compared with the face data.

Before feature extraction all images should be pre-processed and normalized.

As part of pre-processing is dimension reduction of all input images to defined size. Also *contrast limited adaptive histogram equalization* can be applied (**CLAHE**). The normalized images can be masked to omit the background and leave only the face region.

The main objective of the normalization process is to minimize the uncontrolled variations that occur during the acquisition process and to maintain the variations observed in facial feature differences between individuals.

What can bring differences into images is also the pose change.

Feature extraction involves reducing the amount of resources required to describe a large set of data. During face recognition, analysis of big amount of data is performed. Analysis with a large number of variables generally requires a large amount of memory and computation power. Feature extraction is related to reduction of variables and data.

For facial feature extraction the edge detection methods are used most frequently. Very good results are achieved also by *local binary patterns* (**LBP**).

Edge detection is the name for a set of mathematical methods where main aim is to detect points in digital image where brightness changes sharply. These image points with crisply change of brightness are typically organized into a set of curved line segments named edges.

The most frequently used functions for edge detections are Sobel operator (called also Sobel filter), Prewitt operator or Gabor filters.

Extracts features from pre-processed faces can be done via LBP histograms as features. LBP histograms are considered as one of the best features for recognizing faces even when only a limited number of samples is available and can be easily computed in the real time [5] (Fig. 2.1).

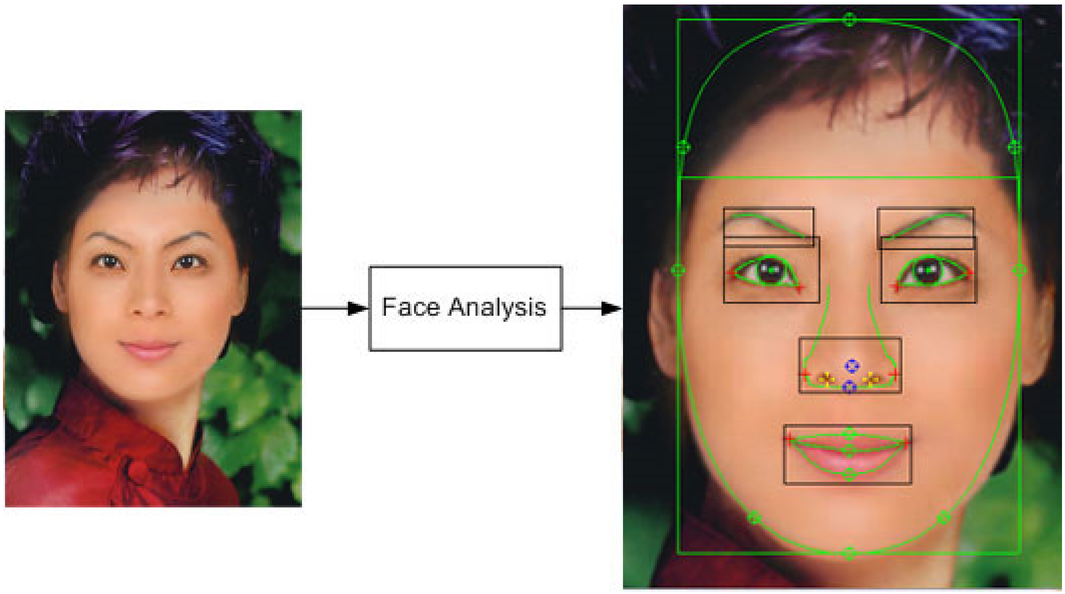


Fig. 2.1 – Example of result for feature extraction

3.3 Faces classification

Face recognition system works usually in two main phases. The first phase is a training process and the second is classification of users. Modern face recognition methods work properly when up to 10 images of one person are available in the training stage. Even the numerous techniques have been developed for face recognition from only a single image per person. Training process should be fully automated and users have to be able to control it. The training process uses clustering algorithms.

The main purpose of all clustering algorithms is to identify clusters or classes in input dataset. There are many clustering algorithms. These algorithms can be divided in two groups: partitioning and hierarchical algorithms [5].

As an example of clustering algorithm K-means can be mentioned. Another algorithm used for clustering is the *self-organizing map* (**SOM**) belonging to neural network techniques or *density-based spatial clustering of applications with noise* (**DBSCAN**).

For the classification of features extracted from faces two methods depending on the number of training images and number of identities which is to be used within the system are listed:

- Support Vector Machines - is used when only relatively small number of identities is considered in the system. Main disadvantage of this method is the time-consuming training of the model when large number of samples is used.

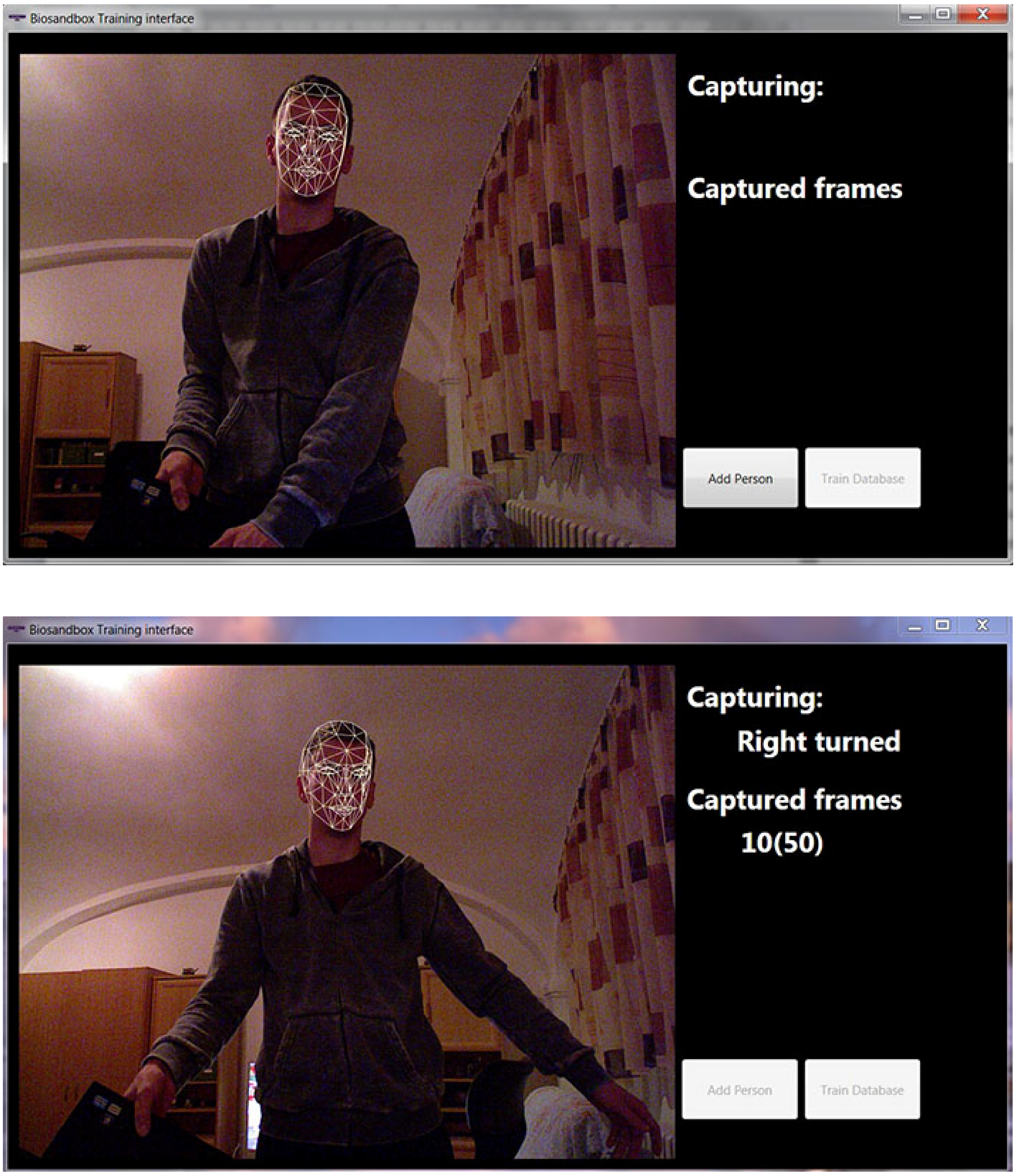
- K-Nearest neighbour distance matching (with the use of Chi-square distance) - this algorithm can be easily parallelized and used in distributed system. The training is done simply by inserting features into the database [5].

3.4 Face localization and recognition

Biometric face recognition systems are widely used in many different types of applications. At present, a smart TV with face recognition system is a typical example of such application. Face recognition in smart TV is used for viewer authentication and based on this, personalized services or different recommendations can be provided. Face recognition systems should works in real time and should be able to recognize one or more identities. The most of this systems include also graphical user interface for automatic training process (Fig. 2.2).

Usually the 2D face recognition task requires processing of the input from a camera. The main face recognition process consists of following sub processes like:

- image acquisition - reads an image from the camera, converts it to the system format and pass it to the system process
- **face localization** - localizes the faces in the image and associate found coordinates with the image. Depending on the camera which is used the localization algorithm is implemented.
- **training process** – clustering algorithms are used , e.g. K-means
- **pre-processing** of localized faces includes histogram equalization
- **normalizatio**n – e.g. resizing
- **feature extraction** - extracts features from pre-processed faces, LBP can be used
- **classification** of faces - use methods like Support Vector Machines or K-Nearest neighbor distance matching
- **face tracking** - usually only frontal faces in the image are tracked because the vast majority of face recognition methods is reliable only with use of frontal face images. Once the face has been recognized, it is tracked, what significantly saves computational resources and can follow the subject even after changes in pose [3]. So the information about recognized user is send as output from the system.



Capturing:

Captured frames

Capturing:
Right turned

Captured frames
10(50)

Fig. 2.2 – Example of training GUI of face recognition system

3.5 Iris recognition

Iris is one of the most popular biometric traits. Combination of touchless scanning, stability over time and high recognition accuracy enable the use in surveillance as well as security applications.

It was shown that the iris recognition accuracy depends on the quality of the captured iris image and image preprocessing. To reduce the negative influence of illumination **NIR** (*near infrarred*) light sensing camera is recommended (Fig. 2.3). Using NIR light allows to add addition light source without influencing comfort of sensing.

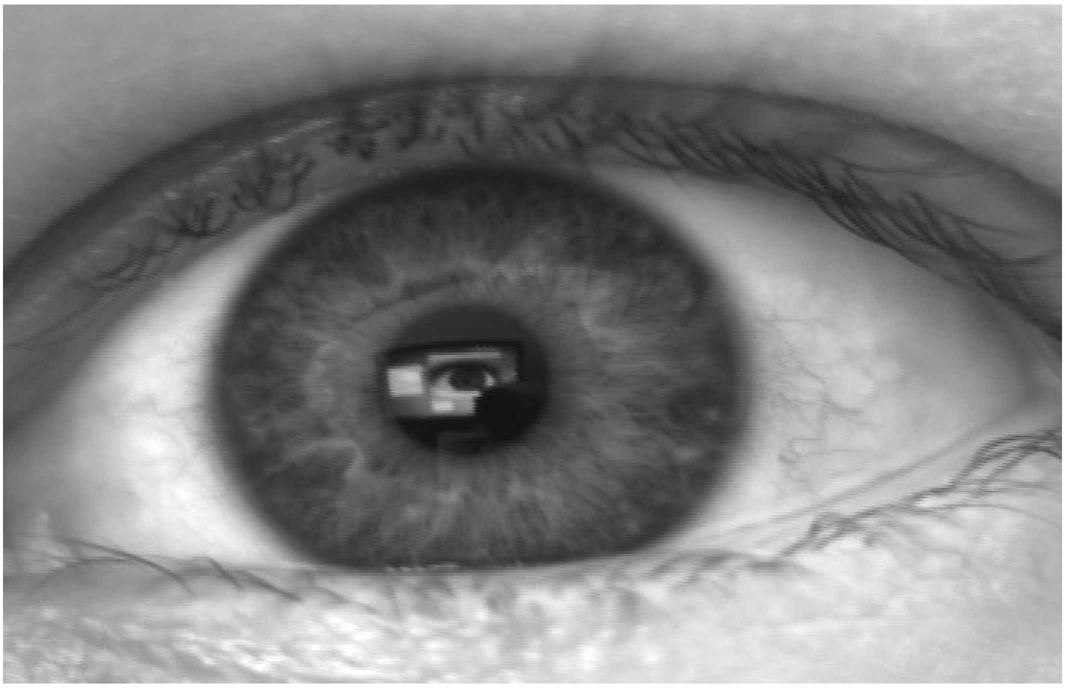


Fig. 2.3 – An example of a picture taken with the Guppy F-038B NIR camera

Iris based identification consists of iris localization, feature extraction, and classification. One of the most successful systems achives 100 percent accuracy in controlled environments. But the localization and normalization for real life application needs to be improved. This system uses Gabor filters for feature extraction where the filtered signals are quantized over two levels. By this procedure strings of binary digits (features) are obtained. By matching the closest samples using the KNN method and a hamming distance recognition is performed.

4 3D face recognition

Face recognition based on 2D face recognition is a common and natural approach. 3D face recognition results in general in higher security than 2D face recognition approach.

3D face recognition based techniques should posses several proprieties such as robustness with respect to lighting variations as well as position, rotation and scaling of the original model within an absolute reference frame [1].

4.1 3D face recognition methods

3D face recognition in comparison to 2D face recognition profits from bigger information flow about the face characteristics. Similarly, both approaches need basic preprocessing like face-size normalization, rotation to a neutral position etc. Added information not only about the 2D shape but about the depth analysis offers rich source of information not captured in 2D images. Main advantages to 2D face analysis are:

- Not affected by illumination variations or use of cosmetics
- Less sensitive to appearance variations
- Easier to handle pose variations
- Projective nature of 2D images
- Simplifies face & facial feature detection, pose estimation & pose compensation

The selected basic 3D face recognition methods:

- Surface-based 3D face recognition - this approach is based on of classic 3D object recognition techniques (Fig. 3.1). There are various types of recognition methods based on:
    - Use of local curvature features, which are rotation invariant (e.g. curve of the face profile)
    - Use of point-to-point matching (polygon of several significant face points)

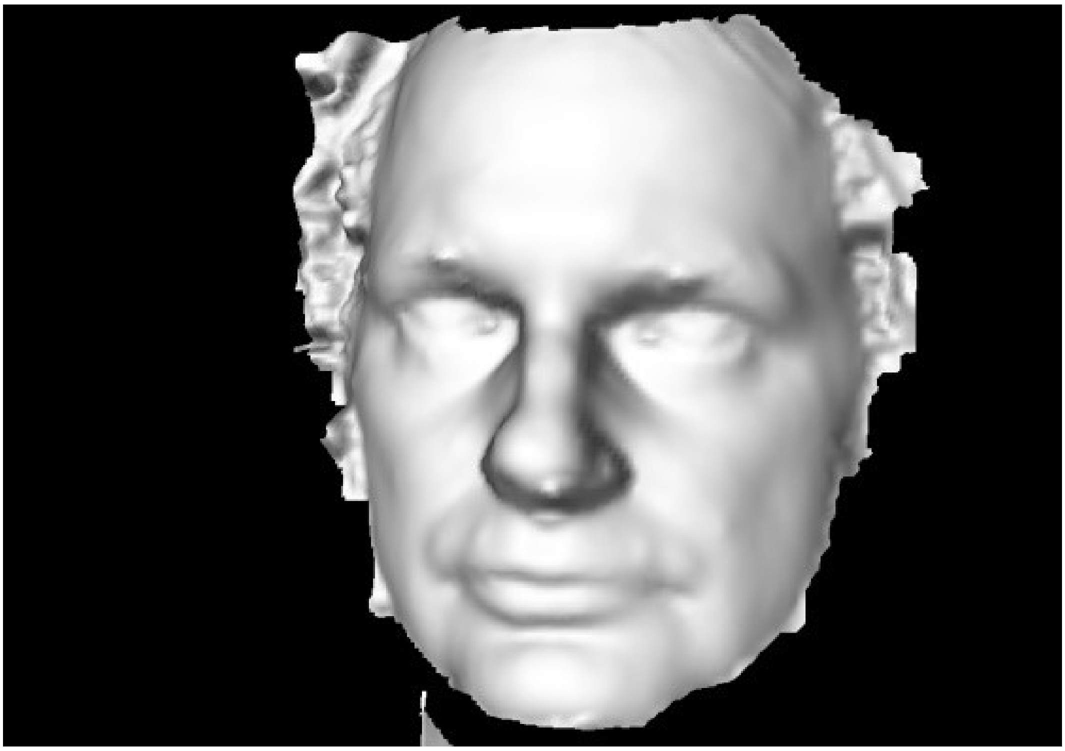


Fig. 3.1 – Surface-based 3D recognition

- Appearance-based 3D face recognition **-** this method deals with eigenfaces and fisherfaces. Accurate alignment of probe and images in database is required. Facial features such as eyes, mouth, etc. are localized and utilized for recognition. The method is easy to implement and is not time consuming (Fig. 3.2).

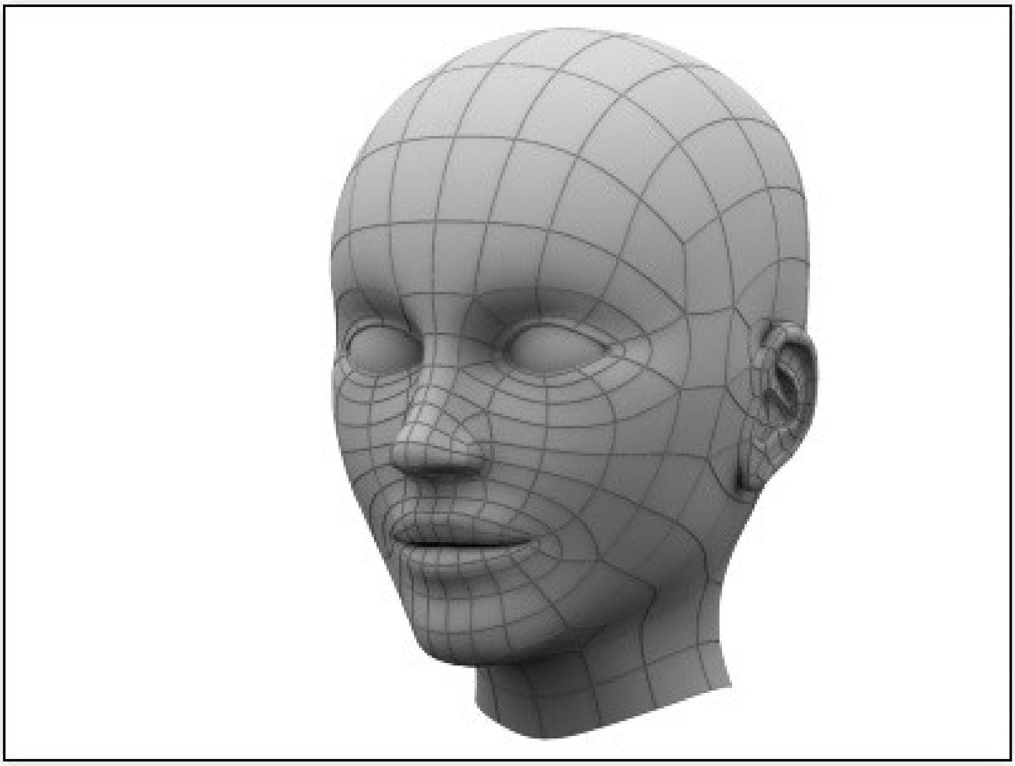


Fig. 3.2 – Appearance-based 3D recognition

- Model- based 3D face recognition **-** this method is based on analysis by synthesis method, where the 3D morphable and annotated face model is produced, which is then compared with models in database. The methods is not suitable for real time application (Fig. 3.3).

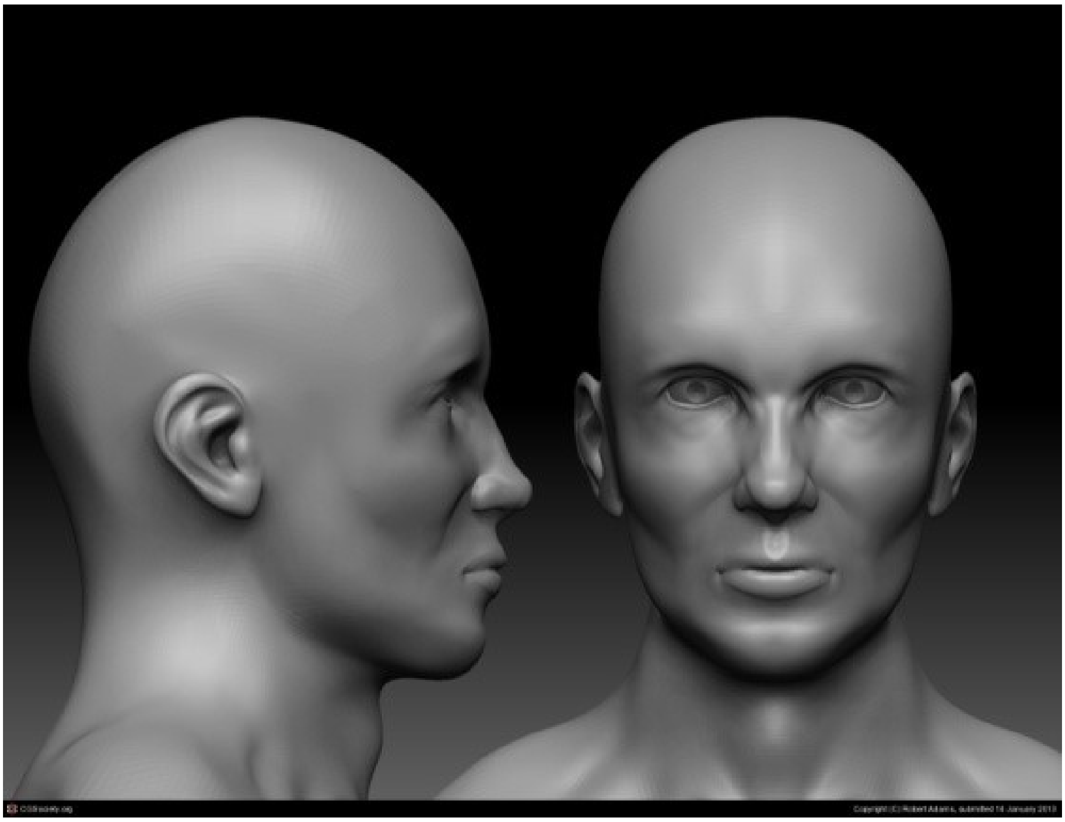


Fig. 3.3 – Model-based 3D recognition

4.2 Preprocessing and data registration

At the beginning of the whole process 3D facial surface is captured (example of creating 3D face is at the pictures Fig. 3.4 – Fig. 3.6). There are several different ways how to achieve this task, for example stereo cameras, depth camera, laser, optical or laser scanner, etc.

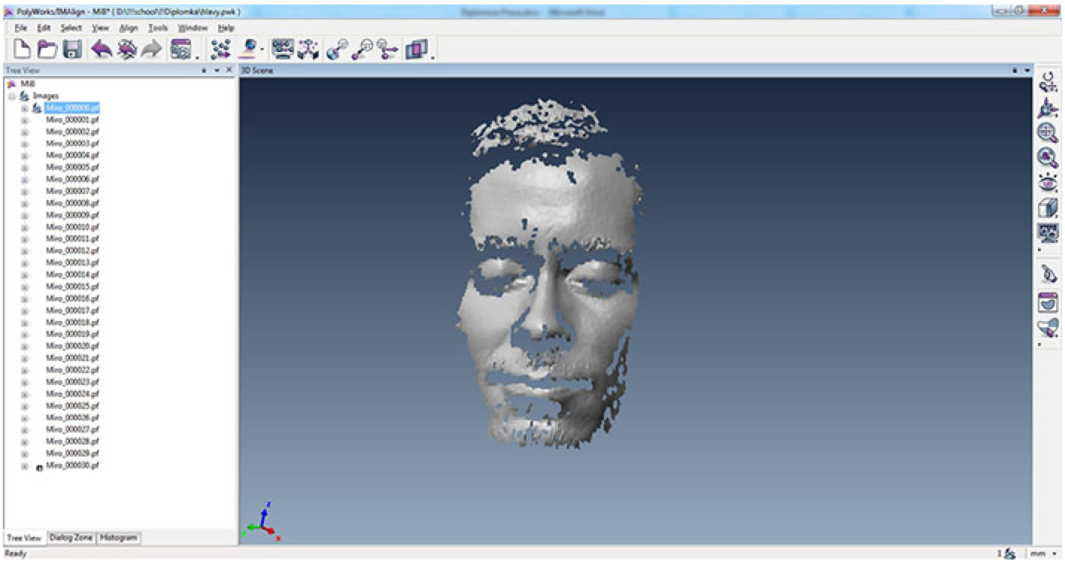


Fig. 3.4 – One scan

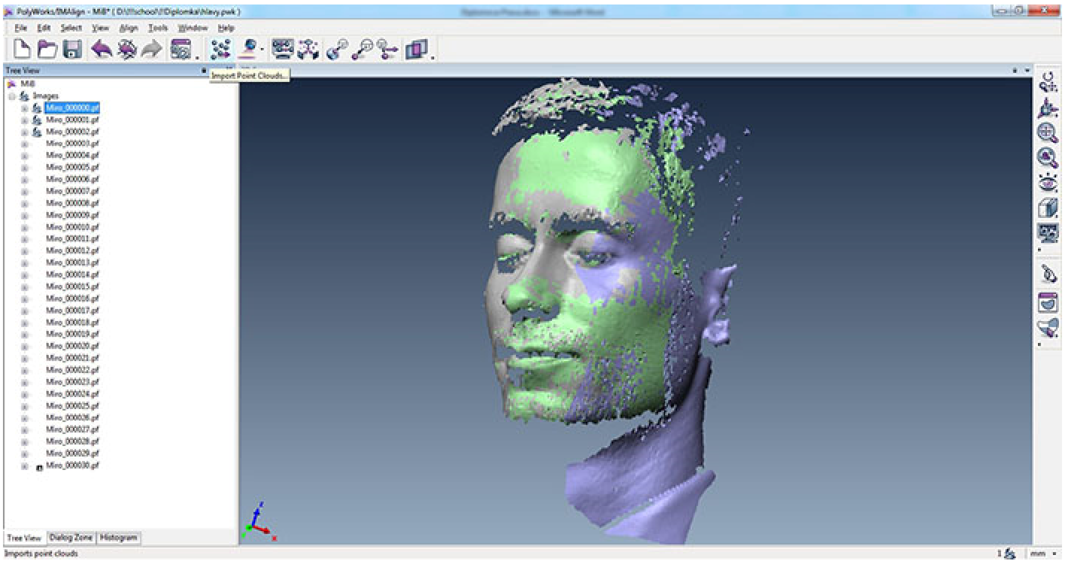


Fig. 3.5 – More scanned pictures create one face

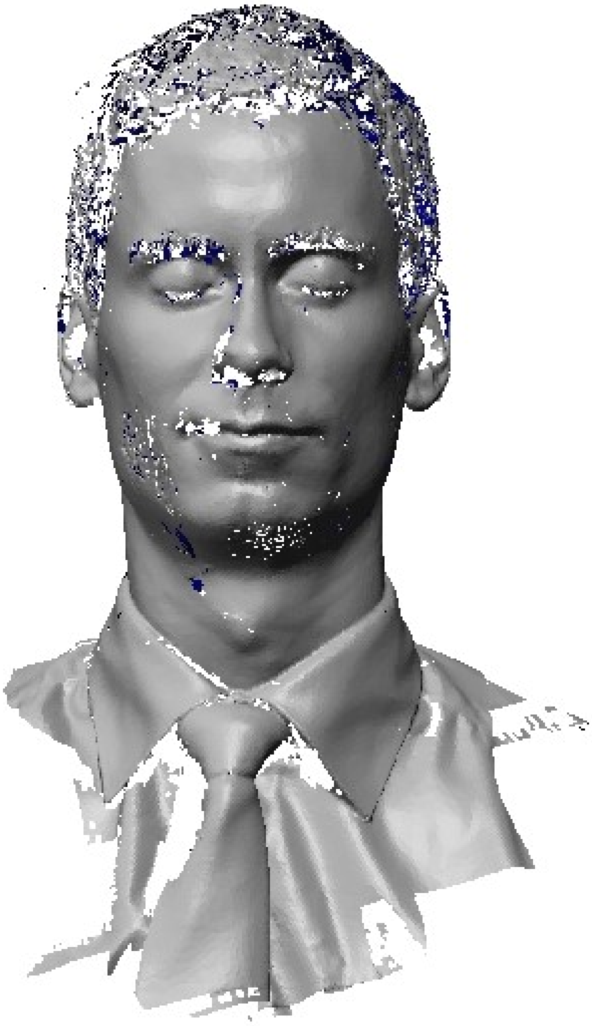


Fig. 3.6 – Final 3D model of face

Only face is needed from whole captured image. Because of that cropping of face is needed. Each face is in the rectangle, which consists of 4 points on the head. The side edges consist from points which position is most left and right. The highest point makes upper edge and the lower edge consist of the lowest point. Then the cropping is based on this rectangle made by these 4 points.

The captured data are subsequently preprocessed using feature extraction algorithms.

The purpose of feature extraction is to extract the compact information from the images that is relevant for distinguishing between the face images of different people and stable in terms of the photometric and geometric variations in the images.

As features can be used facial points (head top, forehead, eyes, chin, nose, mouth, etc.) and distances between these selected points in 3D Euclid space (Fig. 3.7).

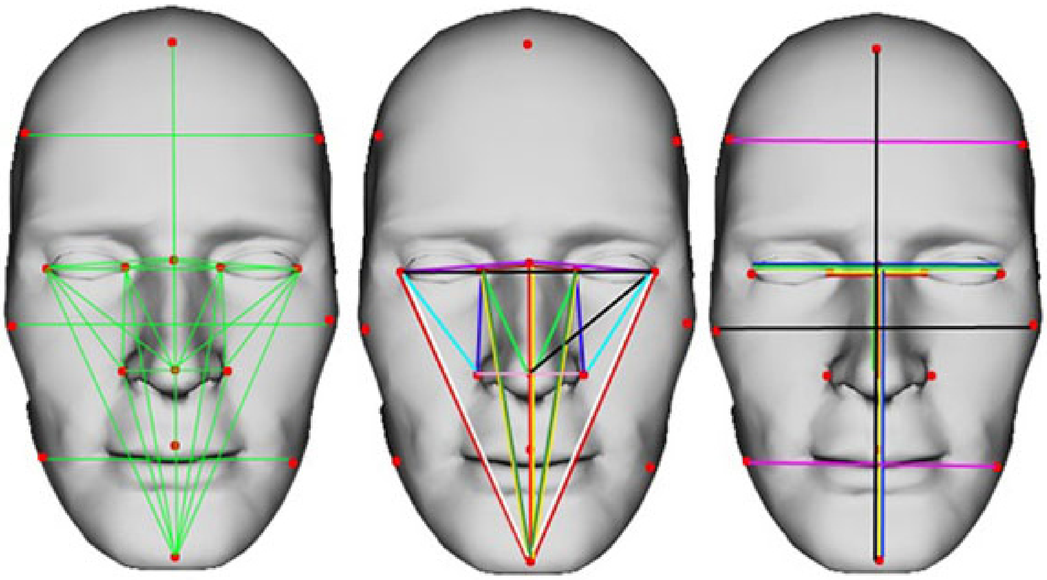


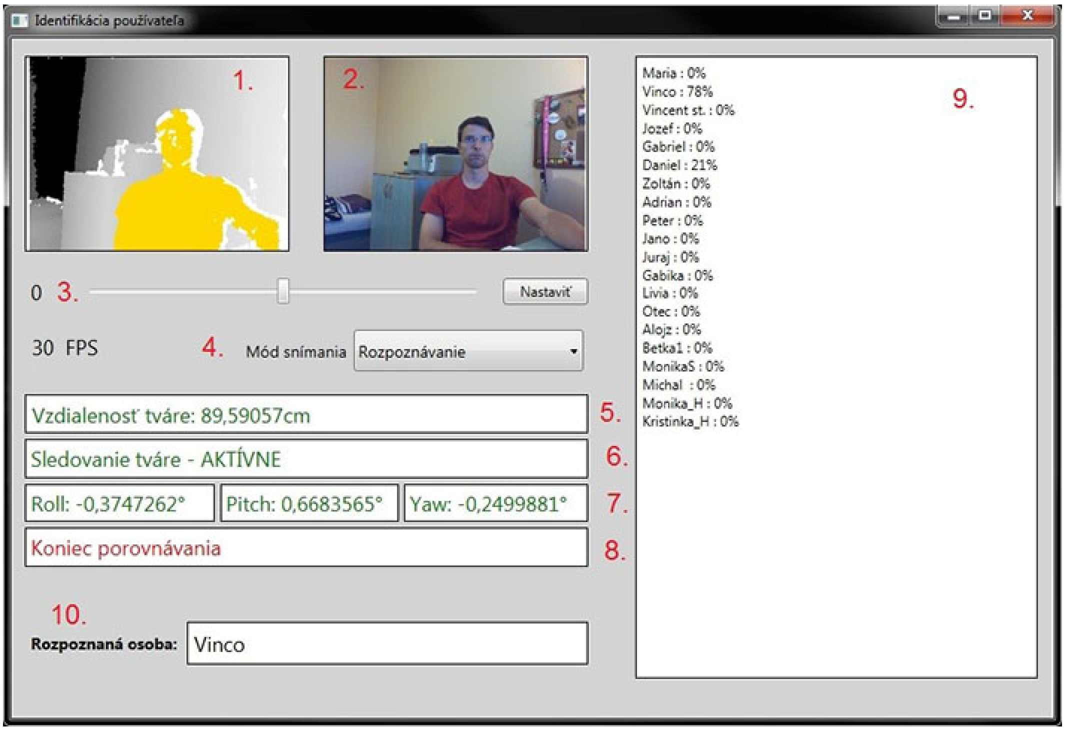
Fig. 3.7 – Example of facial features

4.3 3D face recognition applications

3D face recognition can be also used in many application like secure access into systems or recognize him for smart TV and allow him online shopping (e.g. can be allowed only for parents not for kids, etc.).

The 3D face recognition task requires like 2D face recognition an input from a camera. For 3D face recognition 3D facial surface is needed and has to be captured. The main face recognition process consists of following sub processes like:

- **3D facial surface capturing** - there are several different ways how to achieve this task, for example stereo cameras, laser or depth camera (e.g. of the Kinect sensor), etc.
- **preprocessing** - the captured data are subsequently preprocessed
- **feature extraction** - the purpose of feature extraction is to extract the compact information from the images that is relevant for distinguishing between the face images of different people and stable in terms of the photometric and geometric variations in the images
- **measurement of the distance -**the last step of the 3D face recognition is the measurement of the distance between 3D face of the test user and the 3D faces stored inside of the database. There are several techniques to measure the distance. The simplest method is measuring a local and global distances of two faces where it is needed correctly and very accurate to determine facial points (like eyes, nose, mouth, chin, ears, etc.) and measure their given distances by established metrics. The more sophisticated methods are nearest-neighbor classifier, techniques including support vector machine etc.



30 FPS

Mód snímania: Rozpoznávanie

Vzdialenosť tváre: 89,59057cm

Sledovanie tváre - AKTÍVNE

Roll: -0,3747262° | Pitch: 0,6683565° | Yaw: -0,2499881°

Koniec porovnávania

Rozpoznaná osoba: Vinco

Maria : 0%
Vinco : 78%
Vincent st. : 0%
Jozef : 0%
Gabriel : 0%
Daniel : 21%
Zoltán : 0%
Adrian : 0%
Peter : 0%
Jano : 0%
Juraj : 0%
Gabika : 0%
Livia : 0%
Otec : 0%
Alojz : 0%
Betka1 : 0%
MonikaS : 0%
Michal : 0%
Monika_H : 0%
Kristinka_H : 0%

Fig. 3.8 – Example of GUI for 3D face recognition

5 Authentication

Access security system is proposed with a requirement to allow access only to authorized users whose identity can be verified before. There are essentially three distinct steps, namely the identification, authentication and authorization [6].

Identification – user is identified by token or identification string (phone number or email address)

Authentication – after identification string or token is accepted, user has to prove his identity.

Authorization – Allow or deny user access to the requested content or to a set of actions under based on his access rights.

System can authenticate users based on the assumption the users know something (memometrics), recognize something (cognometrics), own something or has what is characteristic for each person (biometrics). In all three forms the system and user share a secrets (i.e. authentication key).

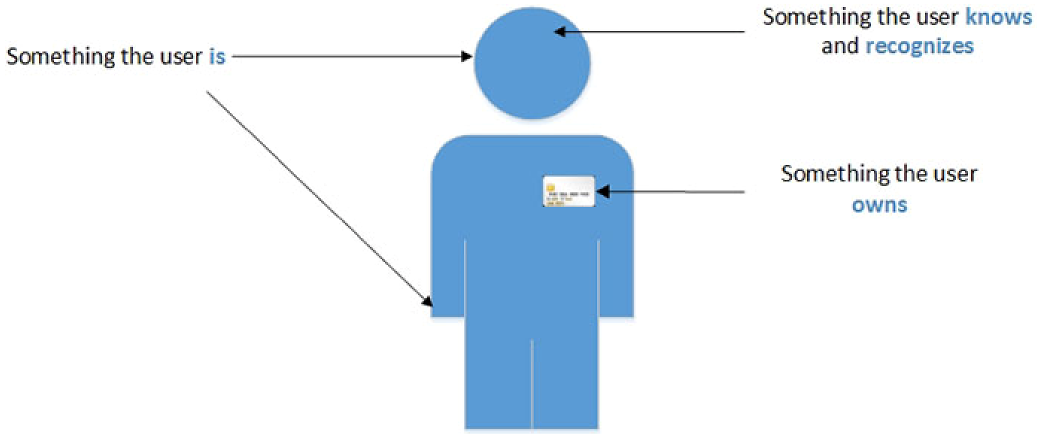


Fig. 4.1 – User authentication options

5.1 Types of authentication mechanisms

Based on listed authentication types the following groups of authentication mechanisms can be listed.

Biometrics

Biometrics is the comparison of anatomical, physiological and behavioral characteristics of a person. Biometric authentication mechanisms fall into two basic categories:

- **Behavioral biometric**s - based on the movements, e.g. the user handling the computer mouse, latency, or the dynamics of keystrokes or signature dynamics.
- **Physiological characteristics** - based on fingerprints, voice, pupil, feature characteristics of face, hand or finger geometry or even the shape of user's ear.

It's difficult to compare biometric technologies within each other. Each has a different range of accuracy, reliability and usability. In case of usability the simple biometric is face detection. Conversely methods that require specific position of the body to the sensor (iris detection), and are thus less comfortable to use, can achieve more accurate results.

Memometrics

This type of authentication mechanism is based on generating random sequences of letters or numbers, called password in case is word or the PIN if it is a numerical expression or passphrase if it contains more than one word. Passwords can be also in semantic form.

Password types:

- random password – the most popular type of authentication with high level of security [7].
- semantic password – are based on deductive process, user is asked several questions with aim to get accurate answer (Fig. 4.2)[8].

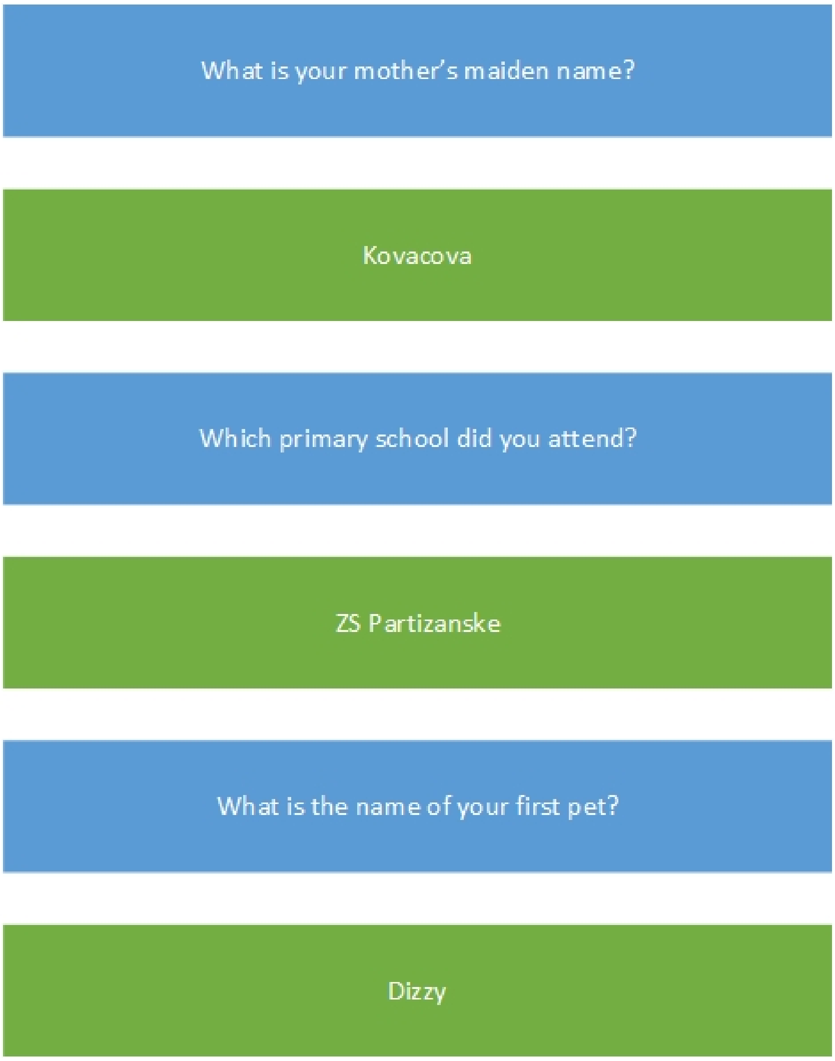

What is your mother's maiden name?
Kovacova
Which primary school did you attend?
ZS Partizanske
What is the name of your first pet?
Dizzy

Fig. 4.2 – Basic principle of semantic password

Cognometrics

Idea graphical authentication is based on the user's visual memory. Scientific studies point to the fact that the human being has a huge and practically unlimited possibilities to remember the pictures [9].

Graphic codes are gaining popularity especially in the case of mobile or tablets technologies, e.g. to unlock mobile phone. There are two main principles:

- **graphic codes based on recognition** - the user selects the target image between the amount of disturbing elements in the scene. This approach is purely based on visual memory. The aim is to recognize previously seen object between the amounts of the other.

- **graphic codes based on the position** - user with this principle must draw a pattern, usually in the grid, which requires visual-spatial memory and precise movement.

Ownership

Authentication can be based on something that a user owns. This object is token. A good example is the token SecureID from RSA Security in Fig. 4.3. [15]

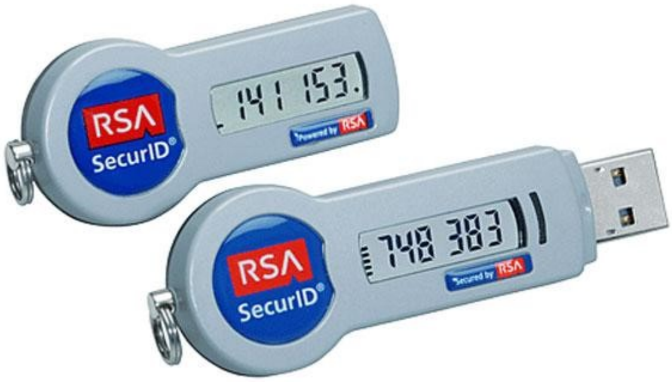


Fig. 4.3 – Token example: SecureID – RSA Security

$E=m\cdot c^2$

Token through a cryptographic function that combines the lock and a secret key, create a numerical code displayed on the LCD. To authenticate user type number from SecureID. The authentication server also knows the secret key stored in the user's token, as well as the time and date. Based on this knowledge the authentication server performs the same cryptographic functions. For successful authentication, the generated value must match the value that was inserted by the user.

Another type of authentication token is the one with **USB** (*Universal Serial Bus*) interface.

Tokens are provided as *software* (**SW**) or *hardware* (**HW**).

–

The main disadvantage of HW token is that user has to always carry it.

The SW tokens are stored in users PC or laptop. In this case user can access the system only from PC where the token is stored.

5.2 Human factors in authentication process

Several authentication scenarios use public key encryption methods (public key cryptography). For example, a user own smart card, which carries the corresponding public key and a private key. During user authentication process, system sends a random challenge. The user signs the challenge with his private key and sends the result. System verifies the signature with a public key. In this way, the system can verify the user holds the right private key without the need to accept his key. Instead storing the public key in a file on the remote system, smart card can present a challenge and signed public key certificate, signed by a third party. This is called *Public Key Infrastructure* (**PKI**) standard and is based on the ITU-T specifications.

Fig. 4.4 shows the entities involved in the authentication process. At each step of this process, a potential attacker can gain access to the authentication key.

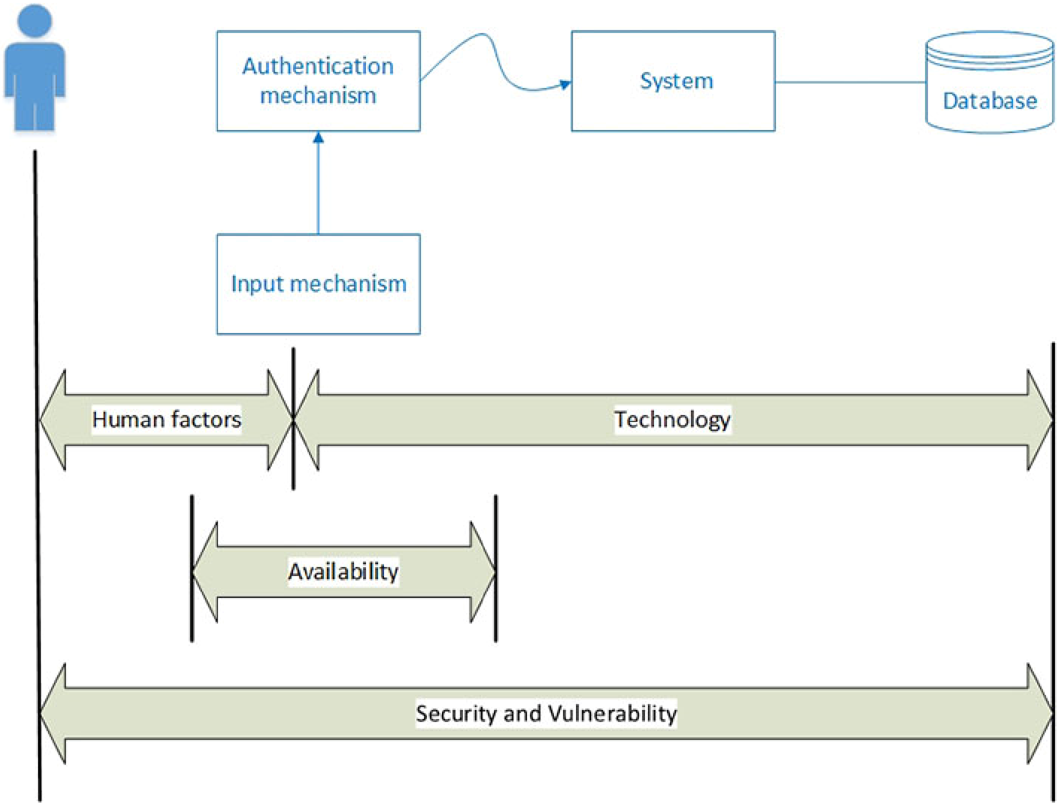


Fig. 4.4 – Entities involved in the authentication process

The most fragile area is the input device and the user. If authentication is based on knowledge (passwords, PIN, etc.) the user has to remember the secret key. Remember password is difficult for many people, they often consciously share their password with someone or write it on paper in the office.

Security can not be solved only with hardware, because users are a part of authentication process [10].

6 Authorization

Authorization means verification of the person at the entrance (to the network or service), based on access rights. In addition, it defines which information can be accessed and what actions can be performed by identified and authenticated user.

6.1 Authorization model

Authorization models (Fig. 5.1) are used to control access rules to the system (or object) and its services, defined by security system. Basic authorization models are [11]:

- *Discretionary Access Control* (**DAC**) – allows the system (or object) owner define who can or cannot access the system
- *Mandatory Access Control* (**MAC**) – user access is defined via classifications.
- *Role-Based Access Control* (**RBAC**) – the most frequently used. Users are divided into groups with defined role. User can access system based on the role.
- *Task Based Access Control* (**TBAC**) – in this model is counter for number of user's access to the system. If defined value is reached, the next access is rejected.
- *Attribute based Access Control* (**ABAC**) – to control access users attributes are used.

All mentioned models can be combined.

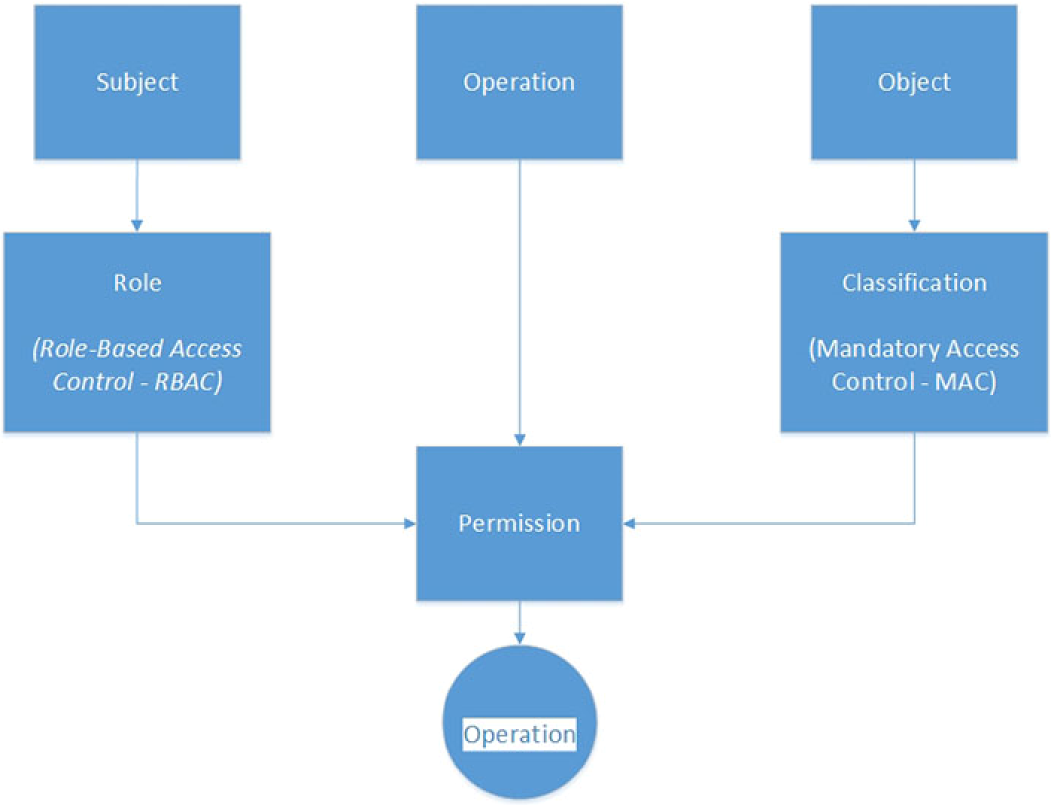

Subject
Operation
Object
Role
(Role-Based Access Control - RBAC)
Classification
(Mandatory Access Control - MAC)
Permission
Operation

Fig. 5.1 – Authorization model

6.2 Access management rules

One of the most common techniques of access control (Fig. 5.2) is access matrix. Rows of the matrix represent user options and columns represent user objects. This technique is often referred as *Access Control List* (**ACL**) [12].

Content-Dependent Access Control is a further extension technique in which one user can access more detailed information or data object as a different user. This decision may depend on factors such as age, used terminal, accesses point, user's access IP address and the time.

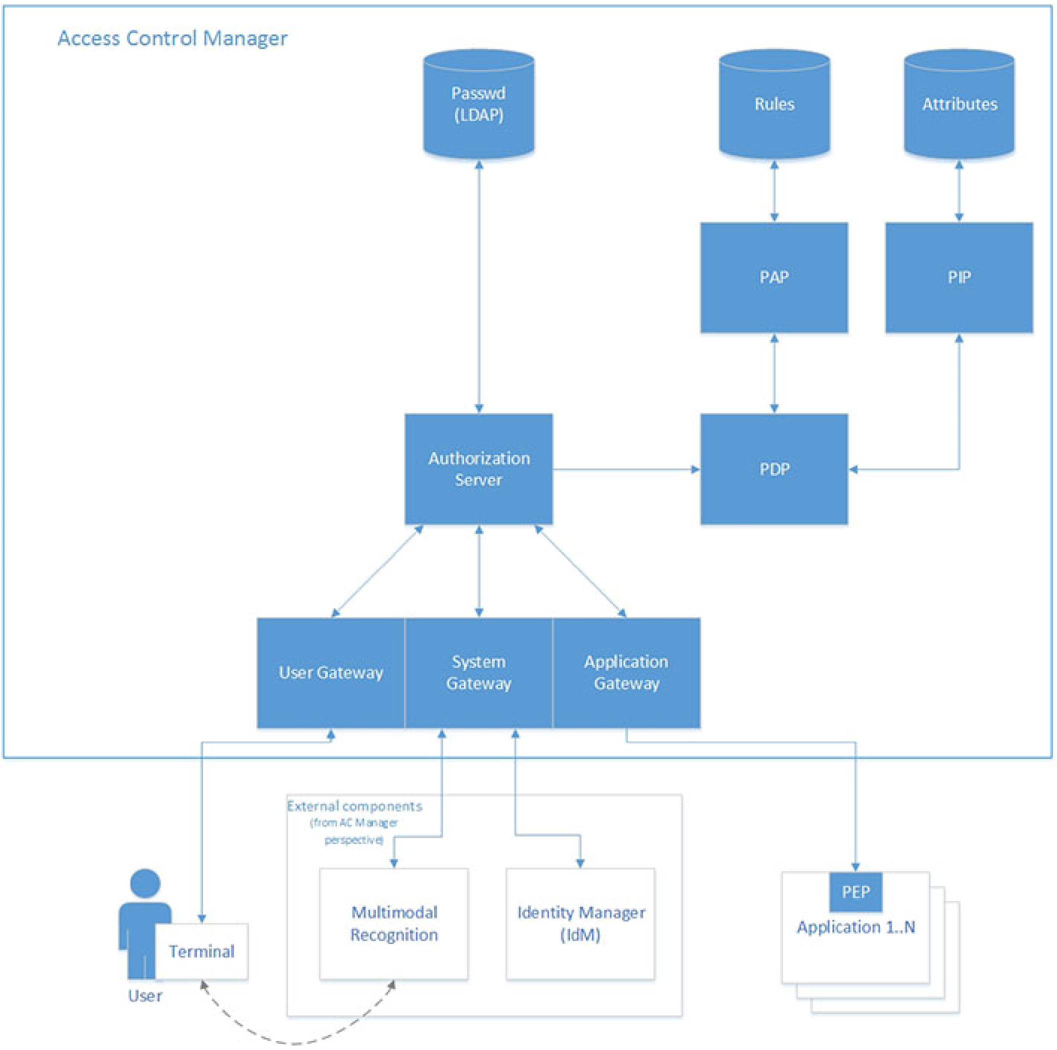


Fig. 5.2 – Access Control Manager

6.3 Access rights

The decision-making process when receiving a request for access to a particular system, application or application information content may, in certain steps, depend on access rights, arranged in an authorization file. Rules allocation is based on the models described in section 5.1 Authorization Model.

Example.

In the system is used RBAC model and are defined three roles:

- administrator
- owner of the group
- user of the group

Administrator assign an owner or users access rights to applications in the system. Group owner can also assign access rights for each user to specific applications in the system. If the administrator has previously allocated to the group owner the rights to add, modify and delete content in a particular application, group owner may further assign those rights to a user. In that case user can also become a contributor of content, i.e. can also act as owner of data. Example of such applications is a service for shared multimedia content.